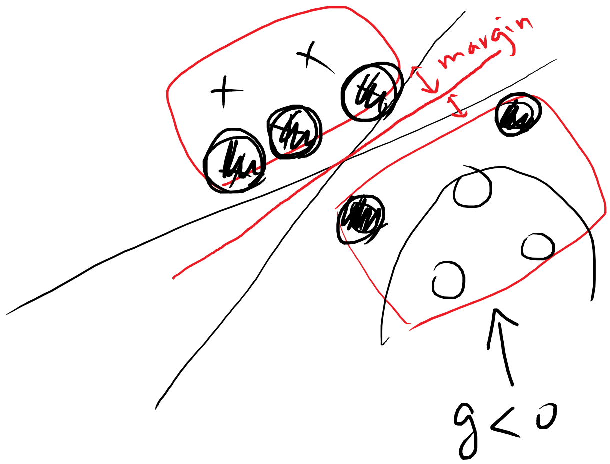
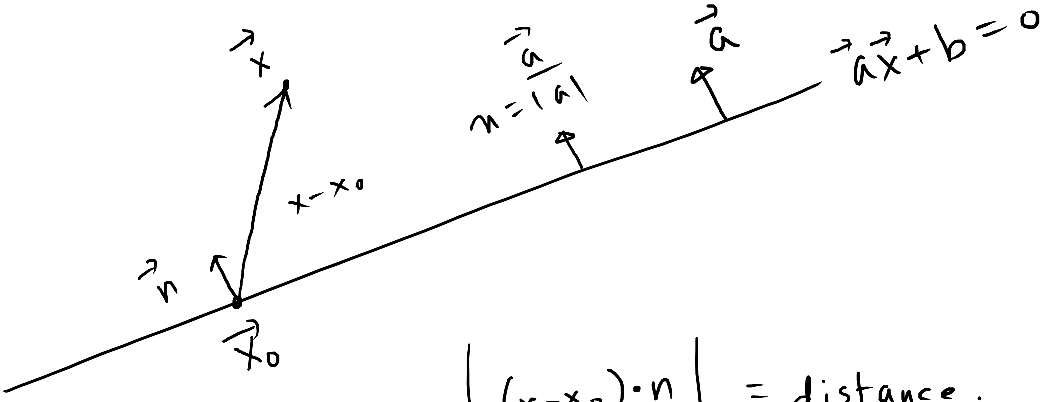
SVM

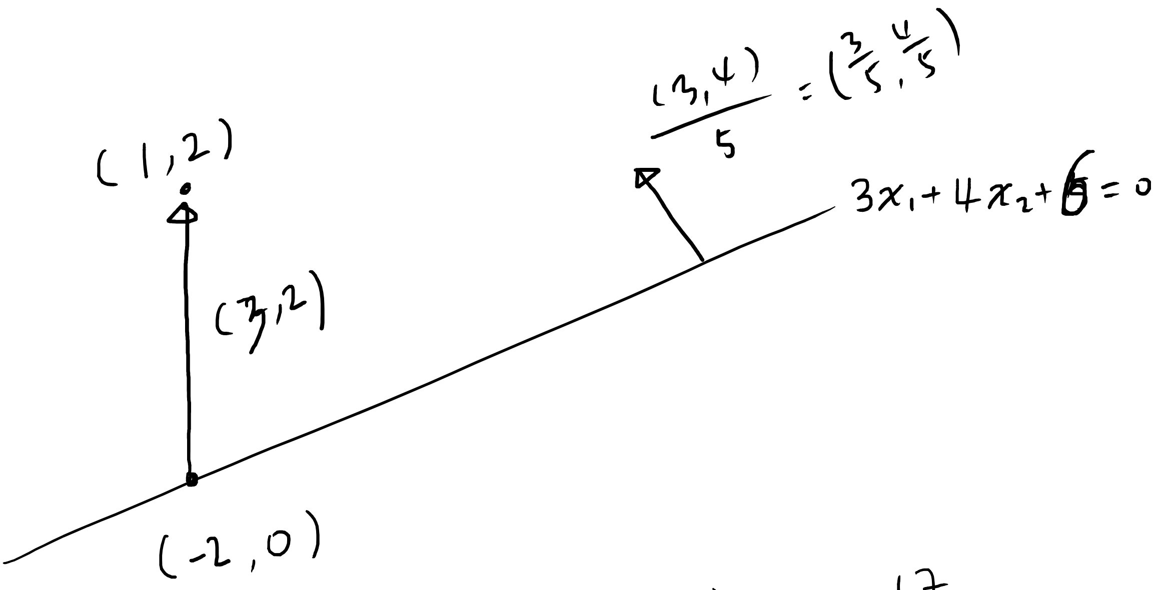
2018년 4월 7일 토요일 오후 1:29



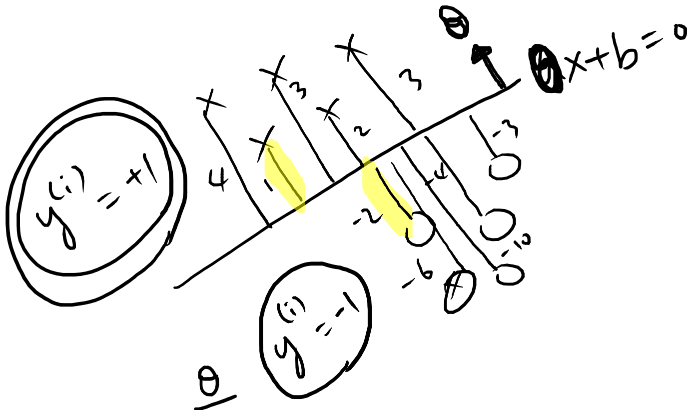
distance between x and P

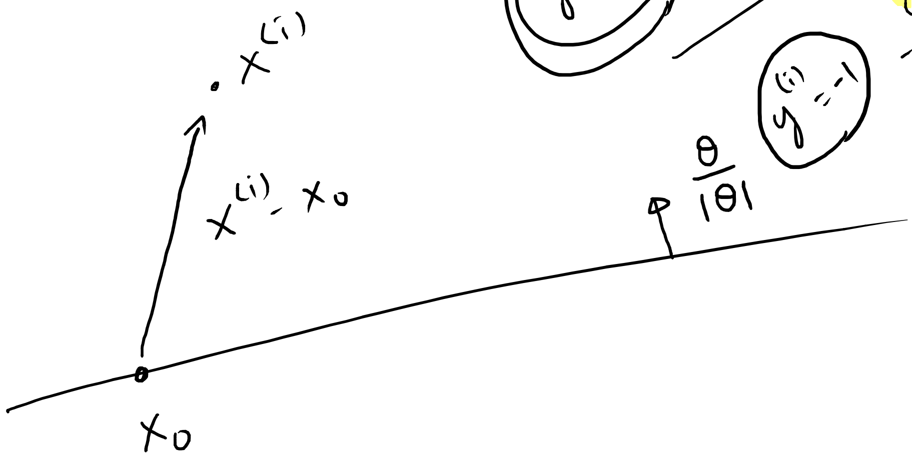


$$\left|(x-x_0)\cdot n\right| = \text{distance.}$$



$$(3,2)\cdot\left(\frac{3}{5},\frac{4}{5}\right) = \frac{17}{5} = 3.4$$





$$\frac{(x^{(i)} - x_0)\cdot\theta}{|\theta|} = \frac{\overset{+b}{\theta x^{(i)}} - \overset{+b \;=\; 0}{\theta x_0}}{|\theta|}$$

$$= \frac{\theta x^{(i)} + b}{|\theta|}$$

$$= \begin{cases} + & \text{if } y^{(i)} = +1 \\ - & \text{if } y^{(i)} = -1 \end{cases}$$

$$\text{margin} = \min_i \; y^{(i)} \left[\frac{\theta x^{(i)} + b}{|\theta|} \right]$$

$$\underset{\theta, b}{\text{maximize}}$$

$$\min_i \; y^{(i)} [\theta x^{(i)} + b] = 0.5$$

$$\Rightarrow \boxed{\theta_*, b_*}$$

$$\min_i \; y^{(i)} [\theta x^{(i)} + b] = 1$$

$$\boxed{2\theta_*, 2b_*}$$

2

convex

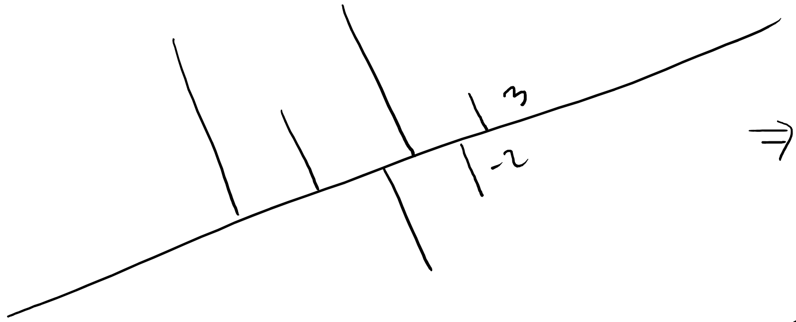
$$\underset{\theta, b}{\text{minimize}} \quad \frac{1}{2}\theta^2$$ convex

$$\boxed{\text{subject to} \quad y^{(i)}\left[\theta x^{(i)} + b\right] \geq 1}$$ convex

$$\boxed{g_i} \leq 0$$

$$= 1 - y^{(i)}\left[\theta x^{(i)} + b\right]$$

"θ"



$\Rightarrow$

$$\text{minimize} \quad \frac{1}{2}\|w\|^2$$

$$\text{subject to} \quad y^{(i)}(wx^{(i)}+b) \geq 1 \quad \text{—} \quad g_i \leq 0$$

$$L = \frac{1}{2}\|w\|^2 + \sum \lambda_i \left[1 - y^{(i)}(wx^{(i)}+b) \right] \quad -\left(\sum_i \lambda_i y^{(i)}\right) b$$

$$w^T w \longrightarrow w + w = 2w$$

$$\boxed{\begin{aligned} Ax &\xrightarrow{\frac{\partial}{\partial x}} A^T \\ x^T A &\xrightarrow{\frac{\partial}{\partial x}} A \end{aligned}}$$

KKT

$$\frac{\partial L}{\partial w} = 0 : \quad \boxed{w + \sum \lambda_i \left[-y^{(i)} x^{(i)} \right] = 0}$$

$$\boxed{w = \sum \lambda_i y^{(i)} x^{(i)}}$$

$$\frac{\partial L}{\partial b} = 0 : \quad \boxed{\sum \lambda_i \left[-y^{(i)} \right] = 0}$$

$$\boxed{\lambda_i \geq 0}$$

$$1 - y^{(i)}(wx^{(i)}+b) \leq 0$$

$$\lambda_i \left[1 - y^{(i)}(wx^{(i)}+b) \right] = 0$$

$$-\sum \lambda_i y^{(i)} w^T x^{(i)}$$

$$x^{(i)T} w$$

$$g(\lambda) = \inf_x \cancel{\text{maximize}}\ L(w, b, \lambda)$$

$$\|w\|^2 = w^T w = \left(\sum_i \lambda_i y^{(i)} x^{(i)T} \right) \left(\sum_j \lambda_j y^{(j)} x^{(j)} \right)$$

$$= \sum_i \sum_j \lambda_i \lambda_j y^{(i)} y^{(j)} x^{(i)T} x^{(j)}$$

$$g(\lambda) = \frac{1}{2} \sum_i \sum_j \lambda_i \lambda_j y^{(i)} y^{(j)} x^{(i)T} x^{(j)} + \sum \lambda_i$$

$$\underset{\lambda}{\text{maximize}}\ g(\lambda) = \frac{1}{2}\sum_i\sum_j \lambda_i\lambda_j y^{(i)} y^{(j)} x^{(i)T} x^{(j)} + \sum \lambda_i$$

$$- \sum_i \lambda_i y^{(i)} x^{(i)T} \sum_j \lambda_j y^{(j)} x^{(j)} + 0$$

$$= \sum \lambda_i - \frac{1}{2}\sum_i\sum_j \lambda_i\lambda_j y^{(i)} y^{(j)} x^{(i)T} x^{(j)}$$

subject to $\lambda_i \geq 0$

$$\sum \lambda_i y^{(i)} = 0$$

	Primal	Dual
변수	w, b	λ_i
# 변수.	$\#w + 1$ (小)	n (大)

In practice.

almost all $\lambda_i = 0$

$\boxed{\text{Dual solved. } \lambda_i \text{ fixed.}}$

$\lambda_i = 0$

$\lambda_i > 0 \rightarrow$ support vector

new data.

$$x \rightarrow \operatorname{sgn}\left[w x + \underline{b} \right]$$

support vector로 계산

$$\sum \lambda_i y^{(i)} x^{(i)} = \sum \lambda_i y^{(i)} x^{(i)}$$

$$\sum \lambda_i y^{(i)} x^{(i)} = \underbrace{\sum_{\substack{\text{support vector} \\ (\lambda_i \neq 0)}} \lambda_i y x}$$

Support vector로 계산

$$\lambda_i \left[1 - y^{(i)} \left(w x^{(i)} + b \right) \right] = 0$$

$i \rightarrow$ support vector $\rightarrow \lambda_i \neq 0$

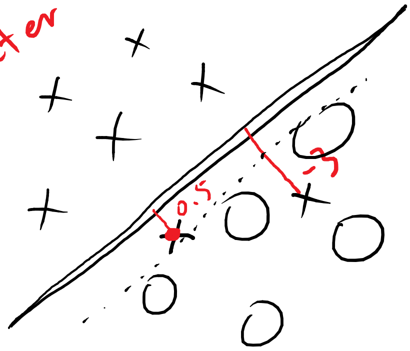
$$1 = y^{(i)} \left(w x^{(i)} + b \right)$$

$$y^{(i)} = w x^{(i)} + b$$

$$b = y^{(i)} - w x^{(i)}$$

$$b = \frac{1}{\#\text{ support vector}} \sum_{\substack{\text{support vector} \\ (\lambda_i \neq 0)}} y^{(i)} - w x^{(i)}$$

fixed ≥ 0
hyperparameter



minimize $\frac{1}{2}\theta^2 + C\sum \xi_i$

subject $y^{(i)}\left[\theta x^{(i)} + b\right] \geq 1 - \xi_i$

$\xi_i \geq 0$

$(w, b, \xi, \lambda_i, \lambda_{i+n})$

$$L = \frac{1}{2}\|w\|^2 + C\sum \xi_i + \sum \lambda_i\left[1 - \xi_i - y^{(i)}\left[w\cancel{x}^{(i)} + b\right]\right] + \sum \lambda_{n+i}\left[-\xi_i\right]$$

$\frac{\partial L}{\partial w} = 0$: $w + \sum \lambda_i\left[-y^{(i)}\cancel{x}^{(i)}\right] = 0 \rightarrow \boxed{w = \sum \lambda_i y^{(i)} x^{(i)}}$

$\frac{\partial L}{\partial b} = 0$: $\boxed{\sum \lambda_i(-y^{(i)}) = 0}$

$\frac{\partial L}{\partial \xi_i} = 0$: $\boxed{C - \lambda_i - \lambda_{n+i} = 0}$

$\lambda_{n+i} > 0 \rightarrow \xi_i = 0$

$0 \leq \lambda_i < C$

$\lambda_i \geq 0$, $y^{(i)}\left[wx^{(i)} + b\right] \geq 1 - \xi_i$, $\lambda_i\left[1 - \xi_i - y^{(i)}\left[wx^{(i)} + b\right]\right] = 0$

$\lambda_{n+i} \geq 0$, $\xi_i \geq 0$, $\lambda_{n+i}\left[-\xi_i\right] = 0$

$$g = -\frac{1}{2}\sum_i\sum_j \lambda_i\lambda_j y^{(i)}y^{(j)} x^{(i)T}x^{(j)}$$

$$\sum_i \lambda_i$$

$$+ C\sum \xi_i + \sum \lambda_i(-\xi_i) + \sum \lambda_{n+i}(-\xi_i)$$

$$= \sum \xi_i \underbrace{(C - \lambda_i - \lambda_{n+i})}_{=0} = 0$$

subject to

$\lambda_i \geq 0$

$\lambda_{n+i} \geq 0$

$\sum \lambda_i y^{(i)} = 0$

$C = \lambda_i + \lambda_{n+i}$

$\Rightarrow 0 \leq \lambda_i \leq C$

$\boxed{\text{dual solved} \quad \lambda_i \quad \text{fixed}} \Rightarrow$

$\lambda_i = 0$

$0 < \lambda_i < C$

$\lambda_i = C$

support vector

$w = \sum \lambda_i y^{(i)} x^{(i)}$ support vector가 결정

support vector

support vector $\lambda_i > 0$

$$\lambda_i \left[1 - \xi_i - y^{(i)}[w x^{(i)} + b] \right] = 0$$

$$y^{(i)}[w x^{(i)} + b] = 1 - \xi_i$$

$\lambda_i < C$

for C

$$\cancel{y^{(i)}}[w x^{(i)} + b] = y^{(i)}$$

for $0 < \lambda_i < C$

$\cancel{y^{(i)}} [w x^{(i)} + b] = y^{(i)}$

$b = y^{(i)} - w x^{(i)}$

$$b = \frac{1}{\# \text{ summation } 0 < \lambda_i < C} \sum_{0 < \lambda_i < C} y^{(i)} - w x^{(i)}$$

new data
x $\longrightarrow$ $\text{sgn} [w x + b]$

primal | dual

w, b, ξ | λ_i, λ_{n+i}

$\boxed{\text{identify support vector.}}$

$\boxed{\text{kernel trick}}$

Kernel trick

$$\left(x^{(i)}, y^{(i)}\right) \xrightarrow{\phi} \left(\phi(x^{(i)}), y^{(i)}\right)$$

$$g = \sum \lambda_i - \frac{1}{2} \sum \lambda_i \lambda_j y^{(i)} y^{(j)} \cancel{\phi(x^{(i)})^T \phi(x^{(j)})}$$

$$K(x^{(i)}, x^{(j)})$$

kernel

subject $0 \le \lambda_i \le C$

$$\sum \lambda_i y^{(i)} = 0$$

$\cancel{K(x, x')} \leftarrow x \Longrightarrow \phi(x)$

dim=3 | dim = 1,000,000,000,000

$$\left(\vec{x}^{(1)} \cdot \vec{x}^{(2)}\right)^2 = (ax+by)^2 = a^2x^2 + b^2y^2 + 2axby$$

$x_1^{(1)}$, $x_2^{(1)}$

$$= \phi(x,y)^T \phi(a,b)$$

$\cancel{x_1^{(i)}, x_2^{(i)}}$

(x, y)

(a, b)

$$= \phi(x, y) \; \phi(a, b)$$

$$\begin{pmatrix} x^2 \\ y^2 \\ \sqrt{2}xy \end{pmatrix} \cdot \begin{pmatrix} a^2 \\ b^2 \\ \sqrt{2}ab \end{pmatrix}$$

$$(x, y) \xrightarrow{\phi} (x^2, y^2, \sqrt{2}xy)$$